DOT/FAA/AM-03/8

Office of Aerospace Medicine
Washington, DC 20591

Development of an Empirically-Based Index of Aircraft Mix

Elaine M. Pfleiderer
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

May 2003

Final Report

U.S. Department
of Transportation

**Federal Aviation
Administration**

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-03/8	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Development of an Empirically-based Index of Aircraft Mix		5. Report Date May 2003
		6. Performing Organization Code
7. Author(s) Pfleiderer EM		8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered
		14. Sponsoring Agency Code
15. Supplemental Notes Work was accomplished under approved task AM-B-02-HRR-522.		

16. Abstract

The present study is part of an ongoing effort to identify objective predictors of subjective air traffic controller workload. The study begins with a comparison of the salient variables governing en route controllers' perceptions of the performance capabilities of a sample of aircraft and the actual performance of the aircraft in the en route environment. A group of 24 Certified Professional Controllers (CPCs) from Kansas City (*N* = 17) and Boston (*N* = 7) en route centers provided estimates of cruising speed, climb, and descent rates for a sample of 24 aircraft types. A matrix of squared Euclidean distances derived from summary measures (i.e., means of estimated speed, climb, and descent rates) was used to construct a classical multidimensional scaling (CMDS) model representing controllers' perceptions of the performance capabilities of each aircraft type. A second matrix was derived from means of speed, climb, and descent rates for the same 24 aircraft types computed from a sample of live air traffic data collected from the Kansas City and Boston en route centers. This matrix was used to construct a second CMDS model representing actual aircraft performance. Interpretation of the dimensions of the CMDS model of ATC estimates suggested that Dimension 1 was related to engine type, whereas Dimension 2 was primarily associated with aircraft weight class. In the model of SAR data, both engine type and weight class were predominantly associated with Dimension 1. Results are used to develop a measure of aircraft mix (i.e., the mix of aircraft with different performance characteristics) to be added to a suite of controller activity and taskload measures.

17. Key Words Aircraft Mix, Controller Workload, Sector Complexity, Traffic Complexity, Multidimensional Scaling		18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 14	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Development of an Empirically-based Index of Aircraft Mix

Aircraft mix has been proposed as one of the traffic characteristics that contributes to sector complexity in en route air traffic control (Robertson, Grossberg, & Richards, 1979; Federal Aviation Administration [FAA], 1984; Grossberg, 1989; Mogford, Murphy, Roske-Hofstrand, Yastrop, & Guttman, 1994). "Sector complexity" describes static and dynamic characteristics of the air traffic control environment that combine with controller taskload (i.e., the air traffic events to which the controller is exposed) to produce controller workload (i.e., the controllers' reaction to and perceived effort involved in managing these events) (Grossberg, 1989; Manning, Mills, Fox, Pfleiderer, & Mogilka, 2001). As changes are introduced into the air traffic control environment such as the recent implementation of the Display System Replacement (DSR), or the proposed introduction of "free flight" (Radio Technical Commission for Aeronautics [RTCA], 1995) it becomes increasingly important that measures are developed to evaluate the impact of these changes on controller performance. In spite of a growing body of work dedicated to the measurement of workload, taskload, sector complexity, and controller performance (for a list of 162 of these measures, see Hadley, Guttman, & Stringer, 1999) little attention has been focused on quantifying aircraft mix. This is possibly because, until recently, aircraft mix had not been clearly defined.

In the initial studies identifying aircraft mix as a factor contributing to sector complexity, aircraft mix was assumed to refer to problems associated with the disparate performance capabilities of propeller and jet aircraft (e.g., Robertson, Grossberg, & Richards, 1979; Grossberg, 1989). In the preliminary stages of a study conducted by Mogford and co-workers (1994), aircraft mix was defined as the proportion of commercial, private, and military aircraft. The number of aircraft flying Visual Flight Rules (VFR) versus Instrument Flight Rules (IFR) was considered to be a distinct factor. A subject-matter expert who provided detailed factor definitions introduced engine type as an element of aircraft mix. In the final list of 19 sector complexity factors, aircraft mix was defined as "VFR, IFR, props, turboprops, jets, etc" (p.37). Though this definition seems extensive, these verbal representations of aircraft mix might be only marginally related to the parameters necessary for quantification of the construct.

Pfleiderer (2000) conducted an investigation of the salient features of aircraft mix as it relates to aircraft performance characteristics. For this analysis, 30 Certified Professional Controllers (CPCs) from various Air Route Traffic Control Centers (ARTCCs) across the United States provided average speed, climb, and descent rate estimates for a sample of 30 distinct aircraft types. A matrix of squared Euclidean distances derived from summary estimates (i.e., means of speed, climb, and descent) was used to construct a classical multidimensional scaling (CMDS) model of the aircraft. Multiple regression interpretation of the two-dimensional solution revealed that Dimension 1 was related to engine type, whereas Dimension 2 was associated with weight class. The results of the analysis were interpreted as evidence of performance-based prototypes. However, it was also evident from the position of the elements (i.e., aircraft types) in the derived stimulus space that it might be possible to develop a measure of aircraft mix using these two easily-obtained variables.

The present study is a continuation of that investigation (i.e., Pfleiderer, 2000). Phase I was designed to determine whether controllers' perceptions of aircraft performance and the actual recorded performance of aircraft were comparable (i.e., would demonstrate similar dimensionality in repeated CMDS analysis). For this analysis, a matrix of squared Euclidean distances of controller estimates of mean speed, climb, and descent rates for 24 distinct aircraft types was compared with a matrix of mean speed, climb, and descent rates of the same aircraft types calculated from routinely-recorded System Analysis Recording (SAR) data. It was expected that the two dimensions noted in the previous CMDS analysis of controllers' perceptions would be the same as those in the SAR sample, but it was possible that the two matrices might differ with regard to the relative salience and importance of each dimension. Characteristics of the CMDS model of SAR data could be used to confirm, amend, or replace previously-gathered information regarding the salient features of aircraft mix.

Phase II focused on the development of an index of aircraft mix based on the results of the Phase I multidimensional scaling analyses. Because multidimensional scaling translates patterns of responding into patterns of elements in a dimensional space, it should be possible to assign base values to aircraft and then calculate distances representing differences in performance capabilities to compute an aircraft mix index.

Finally, the aircraft mix index was computed for all aircraft present in a particular traffic sample. If the index has sufficient variability and precision, it should be able

to discriminate between low-altitude sectors (i.e., sectors with a high probability of aircraft with disparate performance capabilities) and high-altitude sectors (i.e., sectors with a low probability of aircraft mix due to the relatively lower service ceilings of many piston-driven aircraft). If the aircraft mix index passes the "discriminability" test, future research will be conducted to determine whether or not it adds unique information to an existing suite of Performance and Objective Workload Evaluation Research (POWER) measures (Mills, Pfleiderer, & Manning, 2002). It is possible that the complexity associated with aircraft mix is redundant with other variables. It is also possible that aircraft mix is characteristic of so few sectors so as to be of little use within the larger suite of measures. One thing is certain: Aircraft mix's relative contribution to sector complexity and controller workload cannot be assessed until it has been quantified.

Phase I: Comparison of SAR Data and Controller Estimates

Method

Design and Procedure

Multidimensional scaling refers to a group of descriptive procedures that transform data into mapped elements in one or more spatial dimensions (Kruskal & Wish, 1978). The appropriate data for CMDS analysis are proximities, numbers that indicate the similarity or dissimilarity of a set of objects. Proximities may be obtained directly or derived mathematically from a set of variables. In this application, two matrices of dissimilarity measures were computed: One matrix was based on summary controller estimates of aircraft performance, the other was based on summary measures of aircraft performance derived from SAR data.

Controller Estimate Matrix. This matrix represents a subset of the data used in a previous study (Pfleiderer, 2000) in which 30 Certified Professional Controllers (CPCs) provided estimates of average cruising speed, climb rate, and descent rate for each of 30 distinct aircraft types. In the present study, mean speed, climb, and descent rate estimates were calculated from data provided by 24 of the original 30 controllers for 24 of the original aircraft types. The subset of 24 controllers was selected from the larger sample because these CPCs met currency requirements at the same ARTCCs represented in the SAR sample: Kansas City (N = 17) and Boston (N = 7). The aircraft list was reduced because 6 of the 30 aircraft types did not appear in the Kansas City or Boston airspace during the time sampled. The squared Euclidean distance between vectors of controller summary estimates was computed to create a matrix of distances representing controllers' perceptions of each aircraft's capabilities relative to other aircraft in the sample. For example, if summary estimates for the first and second aircraft were:

	Speed	Climb	Descent
Aircraft 1	1	2	3
Aircraft 2	4	5	6

The squared Euclidean distance for these aircraft would be: $(1\text{-}4)^2 + (2\text{-}5)^2 + (3\text{-}6)^2 = 27$. A table listing summary controller estimates used to compute the matrix of distances is provided in Appendix A. For information about participants' professional experience, detailed descriptions of the materials used to collect estimates, and other points of methodology, see Pfleiderer (2000).

SAR Data Matrix. The information used to construct this matrix was recorded at the Kansas City and Boston centers. The Kansas City sample consisted of 168 hours of continuous SAR data recorded from January 19, 1999 through January 25, 1999. The Boston sample comprised a total of 27 hours of SAR data, recorded on March 16, 1998 from 14:00 to 20:59 ZULU (7 hours); March 17, 1998, from 14:00 to 20:59 ZULU (7 hours); March 19, 1998 from 15:00 to 19:59 ZULU (5 hours); and March 20, 1998, from 15:00 to 22:59 ZULU (8 hours). Raw data were extracted through the use of "log" and "track" reports produced by the Data Analysis Reduction Tool (DART). Within the sample time frame, 7095 flights corresponded to the selected aircraft types. The modal flight duration of these flights was 27 minutes. The modal number of updates (observations) per flight was 283.

Aircraft type was derived from designators (alphanumeric labels that indicate the make and model of an aircraft) that are printed on the flight progress strip (FPS) and appear within the flight plan readout display. The contents of both flight progress strips and flight plan readouts are recorded by the Host system and output in the DART log report.

Mean climb and descent rate estimates were calculated from altitude information recorded in the DART track reports. Climb and descent rate estimates represent the amount of change divided by duration of change for all detected altitude changes converted into feet per minute (fpm) and then averaged across changes for each aircraft (set to missing if no altitude changes were detected). For example, from 8:12:08 to 8:17:02 flight XMPL01 climbed from 26,400 feet to 33,000 feet – a total of 6,600 feet in 4 minutes and 54 seconds (1,454 fpm). From 8:30:00 to 8:35:00, XMPL01 climbed from 33,000 feet to 35,000 feet – a total of 2,000 feet in 5 minutes (400 fpm). The climb rate estimate for XMPL01 would then be 927 fpm (the mean of the two changes.) Mean climb and descent rate estimates were calculated in this

manner for each flight and then averaged across flights for each designator. Of course, not all flights made altitude changes during the time sampled, and so the number of observations used to calculate mean climb and descent rates varied between aircraft designators. Appendix A lists the number of climb and descent rate observations upon which mean climb and descent rates were based.

Mean speed estimates were calculated by first computing the mean of all ground speeds recorded in the DART track report for each flight (distinguished by a unique Aircraft Identifier [AID] Computer Identifier [CID] combination) and then averaging across flights for each designator. The number of updates per flight varied as a function of control time. However, the number of speed estimates used to compute the average speed for each designator is equal to the number of aircraft corresponding to that designator (column *N* in the table in Appendix A). Please note that, unlike the computation of mean climb and descent rates, mean speed calculations did not involve speed changes. The measure simply represents the average speed for each aircraft type based on the average ground speed for all individual aircraft of that type.

Squared Euclidean distances were calculated from mean speed, climb, and descent rates in the same manner as controller observations. Distances in the resulting matrix represented each aircraft's performance relative to other aircraft in the sample.

Variables for Interpretation of CMDS Models. A separate set of variables was collected for the purpose of interpreting the dimensions of the CMDS models. The engine number, engine type, and weight class of each aircraft was obtained from information provided in Appendix A of 7110.65N, the most recent version of *Air Traffic Control* (FAA, 2002).

Custom Software. As mentioned previously, the SAR data matrix was computed from DART log and track information for the 7,095 flights with aircraft designators corresponding to the selected aircraft types. These data were extracted from a considerably larger sample of flights. Consequently, it was impractical to manipulate the raw log and track text reports manually. Therefore, two Visual Basic programs were used to extract and organize information and to compute the summary measures. The National Airspace System (NAS) Data Management System (NDMS) program transforms the information in log and track reports into organized database files that provide efficient storage and access (for a more detailed description of this program and its output, see Mills, Pfleiderer, & Manning, 2002). A second program, Aircraft Mix – Phase I, was specially designed to derive pertinent information (i.e., aircraft type, speed, altitude) for each flight from NDMS output tables, compute summary measures, and print this information to SPSS database files for analysis.

Results and Discussion

Controller estimates of speed, climb, and descent rates were compared with mean speed, climb, and descent rates computed from SAR data using Pearson's product moment correlation coefficient. All associations were statistically significant at the $p < .01$ level. However, the magnitude of the correlation between SAR and ATC speed estimates was slightly higher than the others with an $r = .94$. The comparison of mean ATC climb rate estimates and mean climb rates calculated from SAR data produced an $r = .89$, whereas the association between ATC descent rate estimates and SAR descent rate measures resulted in an $r = .79$.

Classical Multidimensional Scaling Analyses (CMDS)

Dimensionality. SAR data and ATC estimate matrices were submitted to separate, non-metric CMDS analyses. The squared correlation is a measure of fit that describes the relationship between the original distances and the derived stimulus coordinates. Squared correlations of the two-dimensional ($r^2 = .9990$) and three-dimensional ($r^2 = .9989$) models of SAR data, and of the two-dimensional ($r^2 = .9995$) and three-dimensional ($r^2 = .9996$) models of ATC estimates suggest that the proximities were described well in either two or three dimensions.

Kruskal's stress formula 1 is a measures of how well the configuration represents the original data. This measure ranges from zero (best fit) to one (worst fit). Stress values for the two-dimensional (.0164) and three-dimensional (.0171) models of SAR data, and of the two-dimensional (.0123) and three-dimensional (.0116) models of ATC estimates indicated that the configuration fit the data well in either two or three dimensions. However, because stress is not a reliable measure in a degenerate solution, scatterplots describing the relationship between proximities and derived distances (i.e., scatterplot of non-linear fit) and disparities (i.e., transformation scatterplot) were examined for both analyses. None of the scatterplots demonstrated patterns characteristic of a degenerate solution. (For an explanation of the causes and caveats of degenerate solutions, see Kruskal & Wish, 1978, p. 29-30.)

In both analyses, the two- and three-dimensional solutions demonstrated excellent fit. Though the three-dimensional solution demonstrated the best fit for the model of ATC estimates, examination of stimulus coordinates and regression analysis failed to uncover any distinctive,

interpretable features of the third dimension. Additional dimensions are of little use if they fail to contribute to the interpretation of the solution (Kruskal & Wish, 1978). Moreover, the two-dimensional model demonstrated the best fit for the SAR data set. Therefore, the two-dimensional solution was selected for interpretation.

Regression Method of Dimensional Interpretation. The most objective technique available for dimensional interpretation is the regression method in which variables believed to correspond with the stimulus configuration are regressed over coordinates. For this application, engine type was coded according to performance capabilities associated with each engine type, from lowest (piston-driven) to highest (jet propelled). Weight class was also coded into three ordered levels: Small aircraft are those of 41,000 lbs. or less maximum certificated takeoff weight; Large aircraft are those of more than 41,000 lbs. up to 255,000 lbs. maximum certificated takeoff weight; Heavy aircraft are those capable of takeoff weights of more than 255,000 lbs. whether or not they are operating at this weight during a particular phase of flight (FAA, 2002).

According to Kruskal and Wish (1978), two conditions are necessary for satisfactory multiple regression interpretation of a dimension. First, the multiple correlations must be extremely high (correlations in the .90s are recommended, although those in the .70s will suffice). As shown in Tables 1 and 2, only engine type and weight class achieved the recommended degree of association with the dimensions. In general, the two data sets were remarkably similar: Correlations were in the .90's for engine type and in the .80's for weight class. However, the two models differed with respect to the relationship between the dimensions and these criterion variables. Notice that in the model of SAR data, the standardized regression weights of both engine type and weight class are more closely associated with Dimension 1 than with Dimension 2. However, in the configuration derived from ATC estimates, the standardized regression weights of engine type were more closely associated with Dimension 1, whereas weight class was more closely associated with Dimension 2.

Table 1. *Summary of Multiple Regression Analysis Interpretation of the Characteristics of the Two-Dimensional CMDS Model of SAR Data*

Criterion	R	R^2	F	p	β_1	β_2
Engine Type	.94	.89	85.92	.00	.87*	-.25*
Weight Class	.80	.65	19.80	.00	.62*	-.44*
Engine Number	.31	.09	1.65	.22	.20	-.28

* p<.01
β_1 Standardized Regression Weights Dimension 1
β_2 Standardized Regression Weights Dimension 2

Table 2. *Summary of Multiple Regression Analysis Interpretation of the Characteristics of the Two-Dimensional CMDS Model of Controller Estimates*

Criterion	R	R^2	F	p	β_1	β_2
Engine Type	.94	.88	80.34	.00	.80*	-.48*
Weight Class	.86	.74	29.16	.00	.48*	-.70*
Engine Number	.43	.18	2.37	.12	.08	.-42

* p<.01
β_1 Standardized Regression Weights Dimension 1
β_2 Standardized Regression Weights Dimension 2

Neighborhood Interpretation of the CMDS Configurations. Neighborhood interpretation focuses on identifying discrete clusters of elements in the CMDS stimulus configuration. Because neighborhood interpretation capitalizes on small distances, this method can sometimes reveal patterns in the data that are not discernable using multiple regression, which attends mostly to large distances. In this particular application, neighborhood interpretation was considered to be an important supplement to dimensional interpretation because it was highly possible that characteristics of the configuration might correspond with categorical variables unsuitable for multiple regression.

In general, the configurations (Figures 1 and 2) were similar. In both models, Group A consisted primarily of piston-driven aircraft. The exception to this was the C208 (Cessna Caravan), a turboprop that did not perform like other turboprops. (As a point of interest, most of the controllers in the sample misclassified the C208 as a piston-driven aircraft.) Group B consisted entirely of turboprops. In both configurations, all aircraft positioned to the right of the dashed gray line are jets. However, the number of groups differed between the configurations. In the model of ATC estimates, high-performance jets are clearly distinguished from other jet aircraft (Group D in Figure 2). In the SAR data model, jets formed a single, loosely-knit group.

Perhaps the most striking difference between the configurations had to do with weight class. Most of the aircraft types in the top portion of the stimulus configuration of ATC estimates (Figure 2) are of the Small and Large weight classes: Heavy aircraft are positioned in the bottom portion of the configuration. This reflects the association of weight class with Dimension 2 in the model of ATC estimates. In the stimulus configuration of SAR data (Figure1), Heavy aircraft are scattered throughout the cluster of jets (Group C).

Conclusions

The clusters of aircraft identified in the neighborhood interpretation of the stimulus configurations present the simplest means by which to code aircraft types for the aircraft mix variable. For the most part, these groups were defined by engine type. Though high-performance jets were not clearly distinguished from other jets in the configuration of SAR data, it seems reasonable to classify these aircraft separately in the computation of the aircraft mix index. On the average, the controllers who contributed estimates for the ATC sample had approximately 10 years of experience at their current ARTCCs. The SAR sample represented 195 hours of traffic. Given the concordance of the two matrices in other respects, it is possible that high-performance jets might have emerged as a separate

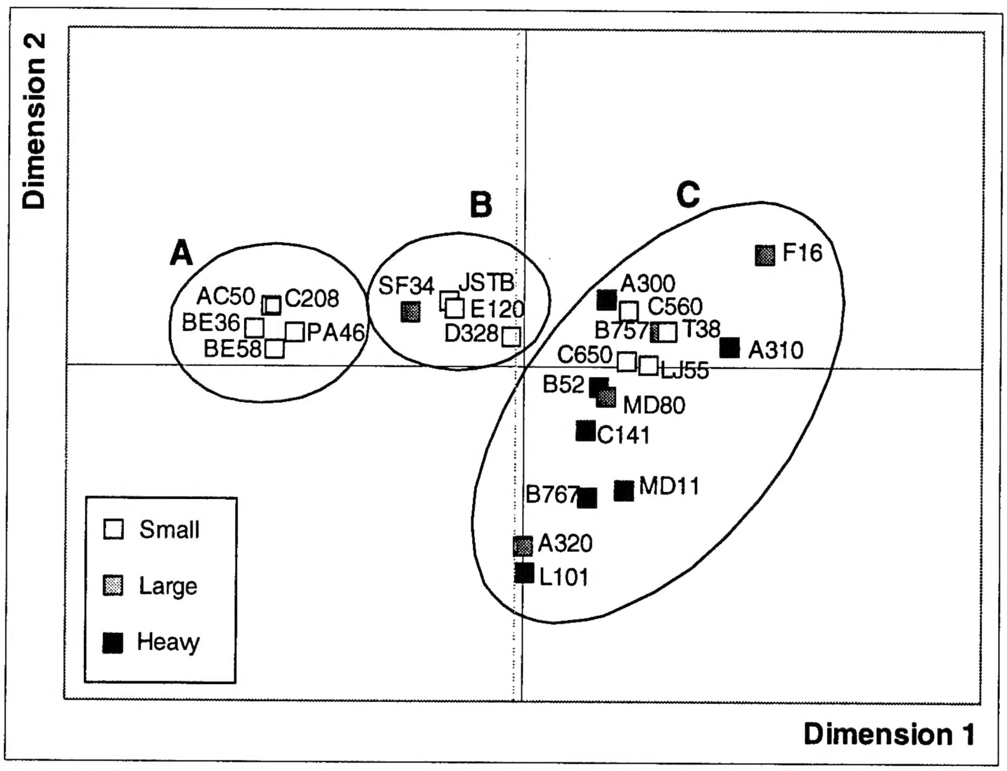

Figure 1. Derived Stimulus Configuration of the Two-Dimensional Classical Multidimensional Scaling (CMDS) Model of SAR Data

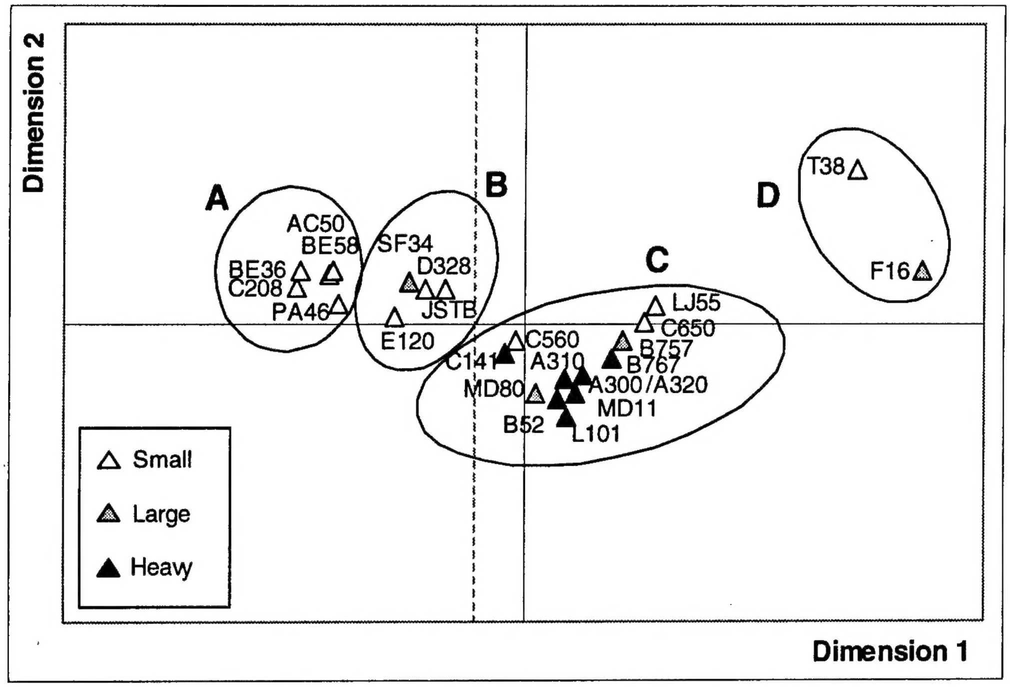

Figure 2. Derived Stimulus Configuration of the Two-Dimensional Classical Multidimensional Scaling (CMDS) Model of ATC Estimates

group in the SAR configuration had the sample been large enough to better approximate the years of experience represented by the controllers in the ATC sample.

It is unlikely that the incorporation of weight class is crucial to the precision of the aircraft mix index. To begin with, weight class is a correlate of engine type (i.e., most piston-driven aircraft are Small, most turboprops are Large, all Heavy aircraft are jets). Because of the nature of this relationship, incorporation of the weight class dimension would only involve jet aircraft (i.e., separating jets into Heavy/other subgroups). However, the tight clustering of the jet aircraft in Group C of the stimulus configuration of ATC estimates (Figure 2) suggests that this differentiation is probably not necessary. Heavy aircraft may perform somewhat differently than other jet aircraft, but this difference appears to be only slightly perceptible to air traffic controllers (other than procedural considerations addressing the wake turbulence associated with Heavy aircraft and B757s).

Phase II: Development and Testing of the Aircraft Mix Index

Method

Sample

The sample selected for testing the aircraft mix index consisted of SAR data from 15 high-altitude sectors and 13 low-altitude sectors within the Kansas City airspace. The Kansas City ARTCC was selected because of the availability of sector information for that particular center (e.g., sector strata, number of underlying airports, sector combinations). The data were recorded on Friday, December 22, 1999 from 15:15 to 16:15 (local time) when most sectors within the Kansas City en route center were open (i.e., sector combinations were minimal).

Procedure

Given the fact that a total of 562 aircraft crossed the Kansas City airspace during the hour sampled, it would be highly impractical to calculate the aircraft mix index manually. Therefore, a Visual Basic program (Aircraft Mix – Phase II) was written to accomplish this task. This program was designed to process SAR information stored in NDMS output tables on a sector-by-sector basis. Based on Phase I results, aircraft were assigned aircraft type codes with values ranging from one to four. Piston-driven aircraft were assigned a value of 1, turboprops a value of 2. With some exceptions, jet aircraft were assigned a value of 3. High-performance jets (i.e., aircraft types that perform within similar parameters as the aircraft in Group D of Figure 2) are coded as such in the system files of all en route Host computers. These file codes were used to assign a value of 4 to all high-performance jets in the sample. Program algorithms were written to produce a half matrix of aircraft type differences between pairs of aircraft within a given sector. Table 3 lists aircraft type differences for the sample of aircraft in Figure 3. For instance, DAL589 is a commercial jet and has been assigned an aircraft mix code

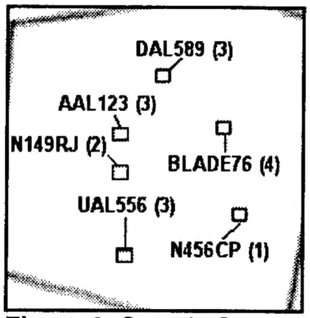

Figure 3. Sample Sector with Aircraft Mix Codes

Table 3. Aircraft Mix Index

	DAL589	AAL123	N149RJ	UAL556	BLADE76
AAL123	0				
N149RJ	1	1			
UAL556	0	0	1		
BLADE76	1	1	2	1	
N456CP	2	2	1	1	3
Index = 17	4	4	4	2	3

Table 4. Descriptive Statistics for Aircraft Mix Index (by Sector Strata)

	Interval 1 (15:15:00-15:29:59)	**Interval 2** (15:30:00-15:44:59)	**Interval 3** (15:45:00-15:59:59)	**Interval 4** (16:00:00-16:14:59)
High Altitude (*N* = 15)				
Mean	1.64	.77	1.28	1.16
S.D.	2.59	1.68	2.29	1.88
Median	.53	.00	.00	.00
Mode	.00	.00	.00	.00
Minimum	.00	.00	.00	.00
Maximum	8.85	5.37	7.63	5.63
Sum	24.58	11.52	19.23	17.36
Low Altitude (*N* = 13)				
Mean	12.35	9.04	8.38	8.95
S.D.	11.90	9.93	6.52	5.53
Median	8.34	4.85	5.48	8.50
Mode	1.50	2.17	3.00	1.11
Minimum	1.50	2.17	3.00	1.11
Maximum	38.93	35.09	24.42	19.70
Sum	160.60	117.46	108.89	116.33

of 3. N149RJ is a turboprop with an aircraft mix code of 2. The aircraft mix difference between N149RJ and DAL589 is 1. The final step in the computation of the index involved summing all items in the half matrix. For example, the aircraft mix index for the group of aircraft in Figure 3 is 17.00 (see Table 3). For each minute of data, the aircraft mix index was calculated for all aircraft pairs at approximately 12-second intervals and stored in an array. At the end of each minute, the mean and standard deviation of the aircraft mix measure were calculated and sent to an array for the purpose of calculating the mean and standard deviation of the aircraft mix measure for each 15-minute interval processed.

Results and Discussion

Computing aircraft mix at 15-minute intervals for high- and low-altitude sectors clearly did not produce a normal distribution (evidenced by the discrepancies between the means and medians of the distributions described in Table 4). For that reason, the Mann-Whitney *U* statistic (Mann & Whitney, 1947) was employed to examine whether the aircraft mix index was reliably different in high- versus low-altitude sectors. The Mann-Whitney *U* is a distribution-free statistic that tests the null hypothesis that two sets of observations were sampled from identical populations. The minimal assumption of the Mann-Whitney *U* is the independence of observations

Table 5. Mann-Whitney *U* Tests for Aircraft Mix Index (by Interval)

	N	Mean Rank	Sum of Ranks	*U*
Interval 1				
High Altitude	15	9.27	139.00	19.00*
Low Altitude	13	20.54	267.00	
Interval 2				
High Altitude	15	9.00	135.00	15.00*
Low Altitude	13	20.85	271.00	
Interval 3				
High Altitude	15	9.13	137.00	17.00*
Low Altitude	13	20.69	269.00	
Interval 4				
High Altitude	15	8.87	133.00	13.00*
Low Altitude	13	21.00	273.00	

* Asymptotic significance (2-tailed) <.01

(Marascuilo & McSweeney, 1977). Given the fact that aircraft cannot be controlled by more than one sector at any given time, it logically follows that aircraft in one sector were independent of aircraft in another within each of the 15-minute intervals.

As shown in Table 5, the null hypothesis was rejected for all comparisons, indicating that the distributions of the aircraft mix index were reliably different in high- versus low-altitude sectors. The sum of ranks assigned to each of the original values for high- and low-altitude sector groups is consistent with the expectation that the mix of aircraft with different performance characteristics (ergo the aircraft mix index) would be higher in low-altitude airspace.

Conclusions

Because the aircraft mix index was able to reliably detect distribution differences in high-and low-altitude sectors, the measure warrants further investigation. Plans for future tests include conducting a CMDS analysis of a much larger sample of aircraft types to include helicopters and other rotorcraft to determine whether they fit into one of the existing aircraft categories or require the introduction of a separate code. Then, the aircraft mix index will be introduced into the current set of Performance and Objective Workload Evaluation Research (POWER) variables (Mills, Pfleiderer, & Manning, 2002) to determine whether or not the aircraft mix index adds unique information to that set. Each step in this process brings us closer to determining the relative contribution of aircraft mix to sector complexity. Constructing the elements that create sector complexity may help us understand the nature of controller workload, and thus provide insight into the relationship between controller workload and performance.

References

Federal Aviation Administration (1984). *Establishment and validation of en route sectors.* (FAA Order No. 7210.46). Washington, DC: Author.

Federal Aviation Administration (2002). *Air traffic control.* (FAA Order No. 7110.65N) Washington, DC: Federal Aviation Administration Air Traffic Operations Program. *Available online*: http://www.faa.gov/atpubs/ATC.

Grossberg, M. (1989). Relation of sector complexity to operational errors. *Quarterly Report of the Federal Aviation Administration's Office of Air Traffic Evaluations and Analysis.* Washington, DC: Federal Aviation Administration.

Hadley, G.A., Guttman, J.A., & Stringer, P.G. (1999, June). *Air traffic control specialist performance measurement database.* (Report No. DOT/FAA/CT-TN95/45).

Kruskal, J.B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills, CA: Sage.

Mann, H.B., & Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*, p. 50-60.

Manning, C.A., Mills, S.H., Fox, C., Pfleiderer, E., & Mogilka, H. (2001). *Investigating the validity of performance and objective workload evaluation research (POWER).* (Report No. DOT/FAA/AM-01/10). Washington, DC: Office of Aerospace Medicine.[1]

Marascuilo, L.A., & McSweeney, M. (1977). *Nonparametric and distribution-free methods for the social sciences.* Monterey, CA: Brooks/Cole.

Mills, S.H., Pfleiderer, E.M., & Manning, C.A. (2002). *POWER: Objective activity and taskload assessment in en route air traffic control.* (Report No. DOT/FAA/AM-02/02). Washington, DC: Office of Aerospace Medicine.[1]

Mogford, R.H., Murphy, E.D., Roske-Hofstrand, R.J., Yastrop, G. & Guttman, J.A. (1994). *Application of research techniques for documenting cognitive processes in air traffic control: Sector complexity and decision making.* (Report No. DOT/FAA/CD-TN94/3). Atlantic City, NJ: Federal Aviation Administration.

Pfleiderer, E.M. (2000). *Multidimensional scaling analysis of controllers' perceptions of aircraft performance characteristics.* (Report No. DOT/FAA/AM-00/24). Washington, DC: Office of Aviation Medicine.[1]

Robertson, A., Grossberg, M., & Richards, J. (1979). *Validation of air traffic controller workload models.* (Report No. DOT/FAA/RD-79/83). Cambridge, MA: U.S. Department of Transportation Research and Special Programs Administration – Volpe National Transportation System Center.

RTCA (1995). *Final Report of RTCA Task Force 3: Free Flight Implementation,* Washington, DC: RTCA Incorporated.

[1]This publication and all Office of Aerospace Medicine technical reports are available in full-text from the Civil Aerospace Medical Institute's publications Web site:

http://www.cami.jccbi.gov/aam-400A/index.html

Copies may be purchased from the National Technical Information Service, Springfield, Virginia 22161.

Appendix A

Summary Estimates/Measures and Number of Observations

Designator	Engine Number	Engine Type	Weight Class		Controller Estimates			SAR Measures		
					Speed (kts)	Climb (fpm)	Descent (fpm)	Speed (kts)	Climb (fpm)	Descent (fpm)
A300	2	J	H	*N*	21	19	19	26	8	19
				Mean	477	2724	2671	364	1730	1380
				SD	34	820	782	79	429	398
A310	2	J	H	*N*	22	20	20	89	51	67
				Mean	466	2458	2653	431	1932	1831
				SD	41	786	869	58	667	509
A320	2	J	L	*N*	23	22	22	614	375	495
				Mean	478	2705	2673	448	1019	1187
				SD	32	706	851	62	531	476
AC50	2	P	S	*N*	23	19	18	151	109	148
				Mean	196	1247	1419	163	856	599
				SD	79	525	550	15	1090	141
B52	8	J	H	*N*	23	22	22	36	20	34
				Mean	476	2386	2498	396	1484	1465
				SD	48	816	799	47	898	694
B757	2	J	L	*N*	23	22	22	158	105	84
				Mean	479	3043	3043	396	1768	1638
				SD	30	1016	1271	64	575	448
B767	2	J	H	*N*	23	22	22	501	227	337
				Mean	488	2898	2923	440	1208	1516
				SD	40	770	1068	65	530	487
BE36	1	P	S	*N*	22	22	22	91	54	80
				Mean	159	1111	1291	165	665	554
				SD	25	457	665	19	250	153
BE58	2	P	S	*N*	23	23	23	240	141	230
				Mean	183	1250	1452	185	701	609
				SD	27	442	683	22	211	187
C141	4	J	H	*N*	22	21	21	19	8	14
				Mean	399	2317	2202	415	1361	1446
				SD	127	889	720	76	632	466
C208	1	T	S	*N*	23	21	20	96	69	90
				Mean	176	1036	1150	161	743	620
				SD	40	289	410	18	214	188
C560	2	J	S	*N*	22	20	20	345	212	248
				Mean	391	2343	2478	356	1684	1658
				SD	93	1025	1176	60	837	431
C650	2	J	S	*N*	23	22	22	234	160	166
				Mean	472	3223	3302	382	1535	1645
				SD	48	1227	1242	68	610	438
D328	2	T	S	*N*	15	13	13	29	17	19
				Mean	275	1873	1888	304	1246	1357
				SD	66	814	818	33	254	226
E120	2	T	S	*N*	22	21	21	95	53	68
				Mean	264	1617	1676	264	1243	1091
				SD	60	747	552	22	263	261

* J = Jet; T = Turboprop; P = Piston
** H = Heavy; L = Large; S = Small

Summary Estimates/Measures and Number of Observations (Continued)

Designator	Engine Number	Engine Type	Weight Class		Controller Estimates Speed (kts)	Controller Estimates Climb (fpm)	Controller Estimates Descent (fpm)	SAR Measures Speed (kts)	SAR Measures Climb (fpm)	SAR Measures Descent (fpm)
F16	1	J	L	*N*	23	19	19	87	68	79
				Mean	585	5379	5316	385	2930	1877
				SD	203	2381	2631	84	3752	645
JSTB	2	T	S	*N*	21	17	18	117	77	81
				Mean	289	2082	2019	254	1218	1089
				SD	76	969	740	35	339	259
L101	3	J	H	*N*	23	23	23	103	62	87
				Mean	497	2439	2513	461	917	1271
				SD	32	718	652	66	380	498
LJ55	2	J	S	*N*	24	20	20	74	41	61
				Mean	463	3303	3498	403	1601	1695
				SD	48	1122	1272	58	652	486
MD11	3	J	H	*N*	22	21	21	66	31	62
				Mean	484	2567	2607	476	1416	1463
				SD	42	643	718	56	537	568
MD80	2	J	L	*N*	23	22	22	3279	1834	2454
				Mean	454	2259	2475	406	1538	1479
				SD	28	615	806	58	1626	376
PA46	1	P	S	*N*	21	21	20	39	24	35
				Mean	199	1314	1485	188	749	691
				SD	37	517	727	42	180	235
SF34	2	T	L	*N*	22	20	19	483	344	394
				Mean	249	1760	1884	241	1087	978
				SD	26	548	787	26	341	237
T38	2	J	S	*N*	24	21	21	123	97	120
				Mean	488	5121	5310	388	2380	1721
				SD	79	2365	2591	66	3435	731

* J = Jet; T = Turboprop; P = Piston
** H = Heavy; L = Large; S = Small